

## **CLAIMS**

1. A method for input ideographic characters comprising the steps of:

(a) entering an input sequence into a user input device;

wherein said user input device comprises:

5           a plurality of input means, each of said input means being associated with a plurality of strokes or phonetic characters, and an input sequence being generated each time when an input is selected by said user input device;

10           data consisting of a plurality of input sequences and, associated with each input sequence, an input method specific database containing a plurality of input sequences and, associated with each input sequence, a set of phonetic sequences whose spellings correspond to the input sequence or a set of strokes sequences corresponding to the input sequence; and

15           an ideographic database containing a set of ideographic character sequences, wherein each ideographic character contains an ideographic index, a plurality of stroke indices to corresponding stroke sequences and a plurality of phonetic indices to corresponding phonetic sequences;

(b) comparing the input sequence with said input method specific database and finding indices to matching strokes entries or phonetic entries and said matching stroke entries or phonetic entries;

20           (c) converting said matching indices to stroke entries or phonetic entries to matching ideographic indices;

(d) retrieving matching ideographic character sequences from said ideographic database by said matching ideographic indices; and

(e) optionally displaying one or more of said matched ideographic character sequences.

2. The method of Claim 1, wherein said stroke indices are indices of strokes sorted by stroke sequences in a stroke input system.

5 3. The method of Claim 2, wherein said stroke input system is a five-stroke or an eight-stroke system.

4. The method of Claim 1, wherein said phonetic indices are indices of phonetic characters sorted by actual spelling in a phonetic input system.

10 5. The method of Claim 4, wherein said phonetic input system is a Pinyin system or a Zhuyin system.

6. The method of Claim 1, wherein said phonetic indices are indices of input means in a phonetic input system.

7. The method of Claim 1, further comprising the step of:

15 prioritizing stroke or phonetic sequences that match an input sequence and prioritizing ideographic character sequences that match a stroke or phonetic sequence according to a linguistic model.

8. The method of Claim 7, wherein said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

20 radical of an ideograph;

radical and number of strokes of a radical;

alphabetical order;

frequency of occurrence of ideographic character sequences, stroke sequences or phonetic sequences in formal, conversational written, or conversational spoken text;

5 frequency of occurrence of ideographic character sequences, stroke sequences or phonetic sequences when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current input sequence entry; and

10 recency of use or repeated use of stroke, phonetic or ideographic character sequences by the user or within an application program.

9. The method of Claim 1, wherein said phonetic sequences comprise single syllables.

10. The method of Claim 1, wherein said phonetic sequences comprise single and multiple syllables.

15 11. The method of Claim 1, wherein said phonetic sequences comprise user generated sequences.

12. The method of Claim 11, wherein in absence of matching phonetic sequences in said database, a sequence of matching phonetic sequences is automatically generated based on single and optionally multiple syllable phonetic sequences.

20 13. The method of Claim 12, wherein said sequence of matching phonetic sequences is narrowed down through user interaction.

14. The method of Claim 12, wherein a sequence of matching ideographic character sequences is automatically generated based on matching phonetic sequences to ideographic character sequences.

15. The method of Claim 14, wherein a sequence of matching ideographic character sequences is narrowed down through user interaction.

16. The method of Claim 7, further comprising the step of:

5 once an ideographic character sequence is selected, changing the associated priority of said matching phonetic sequence and sequence of ideographic characters.

17. The method of Claim 1, wherein the user can specify an explicit ideographic character separator.

18. The method of Claim 1, further comprising the step of:

10 when the user enters a sequence of phonetic characters, returning a sequence of phonetic sequences of exact matches and predictions that partially match.

19. The method of Claim 18, wherein said sequence of phonetic sequences is ordered according to a linguistic model.

15 20. The method of Claim 19, wherein said linguistic model comprises at least one of:

alphabetical order;

frequency of occurrence of phonetic sequences or ideographic character sequences in formal or conversational written text;

20 frequency of occurrence of phonetic sequences or ideographic when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current character sequence entry; and

recency of use or repeated use of phonetic sequences by the user or within an application program.

21. The method of Claim 1, further comprising the step of:

once the user has selected a sequence of ideographic characters, presenting  
5 the user with a list of sequences of one or more ideographic characters.

22. The method of Claim 21, wherein said list of sequences is ordered according to a linguistic model.

23. The method of Claim 22, wherein said linguistic model comprises at least one of:

10 number of total keystrokes in an ideograph;

radical of an ideograph;

radical and number of strokes of radical;

alphabetical order;

frequency of occurrence of ideographic characters in formal or conversational  
15 written text;

frequency of occurrence of ideographic characters when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current character entry; and

20 recency of use or repeated use of ideographic characters by the user or within an application program.

24. The method of Claim 1, wherein the user can enter partial syllables for each of the multiple syllable words.

25. The method of Claim 24, wherein the number of partial keystrokes for each syllable is one.

5 26. The method of Claim 1, wherein one of said plurality of inputs is associated with a special wildcard input that is associated with zero or one of strokes.

27. The method of Claim 1, wherein one of said plurality of inputs is associated with a special wildcard input that is associated with zero or one of said phonetic characters.

10

28. The method of Claim 1, wherein said phonetic indices are indices of phonetic characters sorted by actual spelling in a phonetic input system.

29. A system for receiving input sequences entered by a user and generating textual output in Chinese language, said system comprising:

15 a user input device having a plurality of input means, each of said input means being associated with a plurality of strokes or phonetic characters, an input sequence being generated each time when an input is selected by said user input device;

20 an input method specific database containing a plurality of input sequences and, associated with each input sequence, a set of phonetic sequences whose spellings correspond to the input sequence or a set of strokes sequences corresponding to the input sequence;

25 an ideographic database containing a set of ideographic character sequences, wherein each ideographic character contains an ideographic index, a plurality of stroke indices to corresponding stroke sequences and a plurality of phonetic indices to corresponding phonetic sequences;

means for comparing the input sequence with said input method specific database and finding indices to matching strokes entries or phonetic entries and said matching stroke entries or phonetic entries;

means for converting said matching indices to stroke entries or phonetic  
5 entries to matching ideographic indices;

means for retrieving matching ideographic character sequences from said ideographic database by said matching ideographic indices; and

an output device for displaying one or more matched stroke or phonetic entries, and matched ideographic characters.

10 30. The method of Claim 28, wherein said stroke indices are indices of strokes sorted by stroke sequences in a stroke input system.

31. The system of Claim 29, wherein said stroke input system is 5-stroke or 8-stroke system.

32. The system of Claim 28, wherein said phonetic indices are indices of phonetic  
15 characters sorted by actual spelling in a phonetic input system.

33. The system of Claim 31, wherein said phonetic input system is a Pinyin system or a Zhuyin system.

34. The system of Claim 28, wherein said phonetic indices are indices of input means in a phonetic input system.

20 35. The system of Claim 28, further comprising:

means for prioritizing stroke or phonetic sequences that match an input sequence and prioritizing ideographic character sequences that match a matching stroke or phonetic sequence according to a linguistic model.

36. The system of Claim 34, wherein said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

radical of an ideograph;

5 radical and number of strokes of radical;

alphabetical order;

frequency of occurrence of ideographic character sequences, stroke sequences or phonetic sequences in formal or conversational written text;

10 frequency of occurrence of ideographic character sequences, stroke sequences or phonetic sequences when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current input sequence entry; and

15 recency of use or repeated use of stroke, phonetic or ideographic character sequences by the user or within an application program.

37. The system of Claim 28, wherein said phonetic sequences comprise single syllables.

38. The system of Claim 28, wherein said phonetic sequences comprise both single and multiple syllables.

20 39. The system of Claim 28, wherein said phonetic sequences comprise user generated sequences.



40. The system of Claim 38, wherein in absence of matching phonetic sequences in said database, a sequence of matching phonetic sequences is automatically generated based on single and optionally multiple syllable phonetic sequences.

41. The system of Claim 39, wherein said sequence of matching phonetic sequences is narrowed down through user interaction.

42. The system of Claim 39, wherein a sequence of matching ideographic character sequences is automatically generated based on matching phonetic sequences to ideographic character sequences.

43. The system of Claim 41, wherein a sequence of matching ideographic character sequences is narrowed down through user interaction.

44. The system of Claim 34, further comprising:

means for changing the associated priority of the matching phonetic sequence and the sequence of ideographic characters once an ideographic character sequence is selected.

45. The system of Claim 28, wherein the user can specify a particular tone for the phonetic syllable.

46. The system of Claim 28, wherein one of said plurality of inputs is associated with a special wildcard input that is associated with any or all tones.

47. The system of Claim 28, wherein the user can specify an explicit ideographic character separator.

48. The system of Claim 28, wherein once the user enters a sequence of phonetic characters, the user is returned a sequence of phonetic sequences of exact matches and predictions that partially match.

49. The system of Claim 47, wherein the sequence is ordered according to the frequency of use based on a linguistic model.

50. The system of Claim 48, wherein said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

radical of an ideograph;

5 radical and number of strokes of radical;

alphabetical order;

frequency of occurrence of phonetic sequences or ideographic character sequences in formal or conversational written text;

10 frequency of occurrence of phonetic sequences or ideographic when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current character sequence entry; and

recency of use or repeated use of phonetic sequences by the user or within an application program.

15 51. The system of Claim 28, wherein once the user has selected a sequence of ideographic characters, the user is presented with a list of sequences of one or more ideographic characters.

52. The system of Claim 50, wherein said list of sequences is ordered according to the frequency of use based on a linguistic model.

20 53. The system of Claim 51, where said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

radical of ideograph;

radical and number of strokes of radical;

alphabetical order;

5 frequency of occurrence of ideographic characters in formal or conversational  
written text;

frequency of occurrence of ideographic characters when following a preceding  
character or characters;

grammar of the surrounding sentence;

application context of current character entry; and

10 recency of use or repeated use of ideographic characters by the user or within  
an application program.

54. The system of Claim 28, wherein one of said plurality of inputs is associated  
with a special wildcard input that is associated with zero or one of strokes.

15 55. The system of Claim 28, wherein one of said plurality of inputs is associated  
with a special wildcard input that is associated with zero or one of said phonetic  
characters.

56. A computer usable medium containing instructions in computer readable form  
for carrying out a process for Chinese text entry, said process comprising the steps  
of:

20 (a) entering an input sequence into a user input device;

wherein said user input device comprises:

a plurality of input means, each of said input means being associated with a plurality of strokes or phonetic characters, and an input sequence being generated each time when an input is selected by said user input device;

5           data consisting of a plurality of input sequences and, associated with each input sequence, an input method specific database containing a plurality of input sequences and, associated with each input sequence, a set of phonetic sequences whose spellings correspond to the input sequence or a set of strokes sequences corresponding to the input sequence; and

10           an ideographic database containing a set of ideographic character sequences, wherein each ideographic character contains an ideographic index, a plurality of stroke indices to corresponding stroke sequences and a plurality of phonetic indices to corresponding phonetic sequences;

15           (b) comparing the input sequence with said input method specific database and finding indices to matching strokes entries or phonetic entries and said matching stroke entries or phonetic entries;

          (c) converting said matching indices to stroke entries or phonetic entries to matching ideographic indices;

          (d) retrieving matching ideographic character sequences from said ideographic database by said matching ideographic indices; and

20           (e) optionally displaying one or more of said matched ideographic character sequences.

57.    The medium of Claim 55, wherein said stroke indices are indices of strokes sorted by stroke sequences in a stroke input system.

58.    The medium of Claim 56, wherein said stroke input system is a five-stroke or  
25    an eight-stroke system.

59. The medium of Claim 55, wherein said phonetic indices are indices of phonetic characters sorted by actual spelling in a phonetic input system.

60. The medium of Claim 58, wherein said phonetic input system is a Pinyin system or a Zhuyin system.

5 61. The medium of Claim 55, wherein said phonetic indices are indices of input means in a phonetic input system.

62. The medium of Claim 55, wherein the process further comprises the step of:

10 prioritizing stroke or phonetic sequences that match an input sequence and prioritizing ideographic character sequences that match a stroke or phonetic sequence according to a linguistic model.

63. The medium of Claim 61, wherein said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

radical of an ideograph;

15 radical and number of strokes of a radical;

alphabetical order;

frequency of occurrence of ideographic character sequences, stroke sequences or phonetic sequences in formal, conversational written, or conversational spoken text;

20 frequency of occurrence of ideographic character sequences, stroke sequences or phonetic sequences when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current input sequence entry; and

recency of use or repeated use of stroke, phonetic or ideographic character sequences by the user or within an application program.

5 64. The medium of Claim 55, wherein said phonetic sequences comprise single syllables.

65. The medium of Claim 55, wherein said phonetic sequences comprise single and multiple syllables.

66. The medium of Claim 55, wherein said phonetic sequences comprise user generated sequences.

10 67. The medium of Claim 65, wherein in absence of matching phonetic sequences in said database, a sequence of matching phonetic sequences is automatically generated based on single and optionally multiple syllable phonetic sequences.

15 68. The medium of Claim 66, wherein said sequence of matching phonetic sequences is narrowed down through user interaction.

69. The medium of Claim 66, wherein a sequence of matching ideographic character sequences is automatically generated based on matching phonetic sequences to ideographic character sequences.

20 70. The medium of Claim 68, wherein a sequence of matching ideographic character sequences is narrowed down through user interaction.

71. The medium of Claim 61, wherein the process further comprises the step of:

once an ideographic character sequence is selected, changing the associated priority of said matching phonetic sequence and sequence of ideographic characters.

72. The medium of Claim 55, wherein the user can specify an explicit ideographic character separator.

73. The medium of Claim 55, wherein the process further comprises the step of:

5 when the user enters a sequence of phonetic characters, returning a sequence of phonetic sequences of exact matches and predictions that partially match.

74. The medium of Claim 72, wherein said sequence of phonetic sequences is ordered according to a linguistic model.

10 75. The medium of Claim 73, wherein said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

radical of an ideograph;

radical and number of strokes of radical;

alphabetical order;

15 frequency of occurrence of phonetic sequences or ideographic character sequences in formal or conversational written text;

frequency of occurrence of phonetic sequences or ideographic when following a preceding character or characters;

grammar of the surrounding sentence;

20 application context of current character sequence entry; and

recency of use or repeated use of phonetic sequences by the user or within an application program.

76. The medium of Claim 55, wherein the process further comprises the step of:

once the user has selected a sequence of ideographic characters, presenting the user with a list of sequences of one or more ideographic characters.

77. The medium of Claim 75, wherein said list of sequences is ordered according to a linguistic model.

78. The medium of Claim 76, wherein said linguistic model comprises at least one of:

number of total keystrokes in an ideograph;

radical of an ideograph;

10 radical and number of strokes of radical;

alphabetical order;

frequency of occurrence of ideographic characters in formal or conversational written text;

15 frequency of occurrence of ideographic characters when following a preceding character or characters;

grammar of the surrounding sentence;

application context of current character entry; and

recency of use or repeated use of ideographic characters by the user or within an application program.

20 79. The medium of Claim 55, wherein the user can enter partial syllables for each of the multiple syllable words.



80. The medium of Claim 78, wherein the number of partial keystrokes for each syllable is one.

81. The medium of Claim 55, wherein one of said plurality of inputs is associated with a special wildcard input that is associated with zero or one of strokes.

- 5 82. The medium of Claim 55, wherein one of said plurality of inputs is associated with a special wildcard input that is associated with zero or one of said phonetic characters.